NavWalker: Information Augmented Network Embedding

Kwei-Herng Lai*, Chih-Ming Chen§, Ming-Feng Tsai§, Chuan-Ju Wang*
*Research Center for Information Technology Innovation, Academia Sinica
§Department of Computer Science, National Chengchi University
Email: *{khlai,cjwang}@citi.sinica.edu.tw, §{104761501,mftsai}@nccu.edu.tw
Brief Introduction

Toward understanding network embedding and its applications
“Automatically discover and map a network’s structural properties into a latent space.”

Ref: Network Embedding as Matrix Factorization: Unifying DeepWalk, LINE, PTE, and node2vec, Proceedings of WSDM’18
Applications

Community Detection
• Clustering
  • KNN
• Visualization
  • t-SNE, PCA
Applications

Community Detection
• Clustering
  • KNN
• Visualization
  • t-SNE, PCA
Applications

Community Detection
- Clustering
  - KNN
- Visualization
  - t-SNE, PCA

Node Classification
- Labeling full graph
  - Document Classification
  - Protein Classification
Applications

Community Detection
- Clustering
  - KNN
- Visualization
  - t-SNE, PCA

Node Classification
- Labeling full graph
  - Document Classification
  - Protein Classification

Link Prediction
- Predicting missing relations
  - Recommender system
  - Latent structure prediction
Related Works

[A bank is a financial institution] that accepts deposits from the public and creates credit.

**Distributional Hypothesis**

**Word2vec**

- Distributional Hypothesis
- Skip-gram/CBOW
Related Works

Distributional Hypothesis

Word2vec
- Distributional Hypothesis
- Skip-gram/CBOB

LINE
- 1st/2nd proximity
- Edge sampling

DeepWalk
- Skip-gram
- Random walk

"A bank is a financial institution that accepts deposits from the public and creates credit."
Related Works

Distributional Hypothesis

Word2vec
- Distributional Hypothesis
- Skip-gram/CBOW

LINE
- 1st/2nd proximity
- Edge sampling

DeepWalk
- Skip-gram
- Random walk

HPE
- LINE + DeepWalk

[A bank is a financial institution that accepts] deposits from the public and creates credit.
Related Works

Distributional Hypothesis

Word2vec
- Distributional Hypothesis
- Skip-gram/CBOB

DeepWalk
- Skip-gram
- Random walk

LINE
- 1st/2nd proximity
- Edge sampling

HPE
- LINE + DeepWalk

node2vec
- Alias method (Biased RW)
- Homophily hypothesis (BFS)
- Structural equivalence (DFS)

A bank is a financial institution that accepts deposits from the public and creates credit.
Bank is a financial institution that accepts deposits from the public and creates credit.
“Why should we use auxiliary information?”
NavWalker

Information Augmentation
Information Augmentation - Closeness Centrality

\[ C(i) = \frac{n - 1}{\sum_{v_i, v_j \in V} \text{dist}(i, j)} \]
Information Augmentation - Closeness Centrality

\[ C(i) = \frac{n - 1}{\sum_{v_i, v_j \in V} \text{dist}(i, j)} \]

\[ T = \begin{bmatrix}
  t_{11} & t_{12} & \cdots & t_{1n} \\
  t_{21} & t_{22} & \cdots & t_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  t_{n1} & t_{n2} & \cdots & t_{nn}
\end{bmatrix} \]

1. \[ t_{ij} = C(j) \]

2. \[ t_{ij} = C^{-1}(j) \]
Information Augmentation

\[ A = \begin{bmatrix}
    a_{11} & a_{12} & \cdots & a_{1n} \\
    a_{21} & a_{22} & \cdots & a_{2j} \\
    \vdots & \vdots & \ddots & \vdots \\
    a_{n1} & a_{n2} & \cdots & a_{nn}
\end{bmatrix}, \quad n = |V| \]

\[ T = \begin{bmatrix}
    t_{11} & t_{12} & \cdots & t_{1n} \\
    t_{21} & t_{22} & \cdots & t_{2n} \\
    \vdots & \vdots & \ddots & \vdots \\
    t_{n1} & t_{n2} & \cdots & t_{nn}
\end{bmatrix}, \quad n = |V| \]
Information Augmentation

\[ A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} , \quad n = |V| \]

\[ T = \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ t_{21} & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ t_{n1} & t_{n2} & \cdots & t_{nn} \end{bmatrix} , \quad n = |V| \]

\[ A' = A \circ T \]
Information Augmentation

\[ A' = \begin{bmatrix}
  a'_{11} & a'_{12} & \cdots & a'_{1n} \\
a'_{21} & a'_{22} & \cdots & a'_{2j} \\
  \vdots & \vdots & \ddots & \vdots \\
a'_{n1} & a'_{n2} & \cdots & a'_{nn}
\end{bmatrix}, \quad n = |V| \]

\[ P = \begin{bmatrix}
p_{11} & p_{12} & \cdots & p_{1n} \\
p_{21} & p_{22} & \cdots & p_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
p_{n1} & p_{tn2} & \cdots & p_{nn}
\end{bmatrix}, \quad n = |V| \]

\[ \sum_{j=1}^{n} p_{ij} = 1 \quad \text{for } i = 1, \ldots, n \]
Information Augmentation

\[ O = - \sum_{v_i \in V} \log P(N(v_i) | \Phi(v_i)) \]
“How to use auxiliary information PROPERLY?”
NavWalker
Rooted Random Walk Sampling
Random Walk Sampling

From node A to node G
Neighborhood order: 3

NavWalker

Walk Length = 3
Number of nodes = 5

SkipGram

Window Size = 3
Number of nodes = 5
Random Walk Sampling

From node A to node G
Neighborhood order: 3

NavWalker

Path: 1. A-B-F-J
Walk Length = 3
Number of nodes = 5

SkipGram

Path: 1. A-B-F-J-G
Context Window:
• A: [A,B,F,J]
Window Size = 3
Number of nodes = 5
Random Walk Sampling

From node A to node G
Neighborhood order: 3

Path:
1. A-B-F-J
2. B-F-J-G

NavWalker

Walk Length = 3
Number of nodes = 5

SkipGram

Path:
1. A-B-F-J-G

Context Window:
- A: [A,B,F,J]
- B: [A,B,F,J,G]

Root/Center Node
Neighbor Node

Walk Length = 3
Number of nodes = 5

Window Size = 3
Number of nodes = 5
Random Walk Sampling

NavWalker

From node A to node G
Neighborhood order: 3

Path:
1. A-B-F-J
2. B-F-J-G
3. F-J-G-I

Walk Length = 3
Number of nodes = 5

SkipGram

Path:
1. A-B-F-J-G

Context Window:
• A: [A,B,F,J]
• B: [A,B,F,J,G]
• F: [A,B,F,J,G]

Window Size = 3
Number of nodes = 5
From node A to node G
Neighborhood order: 3

Path:
1. A-B-F-J
2. B-F-J-G
3. F-J-G-I
4. J-G-I-H

Walk Length = 3
Number of nodes = 5

Context Window:
- A: [A,B,F,J]
- B: [A,B,F,J,G]
- F: [A,B,F,J,G]
- J: [A,B,F,J,G]
Random Walk Sampling

From node A to node G
Neighborhood order: 3

Path:
1. A-B-F-J
2. B-F-J-G
3. F-J-G-I
4. J-G-I-H
5. G-I-H-D

Walk Length = 3
Number of nodes = 5

NavWalker

SkipGram

Context Window:
- A: [A,B,F,J]
- B: [A,B,F,J,G]
- F: [A,B,F,J,G]
- J: [A,B,F,J,G]
- G: [B,F,J,G]
Information Augmentation

\[ p(v_j | v_i) = (p_j) \in \left[ \sum_{k=1}^{n} (P)^k \right]_i \]
Experiment
Classification & Recommendation
Experiment — Settings

• Multi-label Classification:
  • Label:
    • Blogger classification
    • Biological state classification
    • POS tag classification
  • Classifier:
    • Logistic regression (L2)

| Classification      | |V|   | |E|   | #(labels) |
|---------------------|---------------------|---------------------|---------------------|
| BlogCatalog         | 10,312              | 333,983             | 39                  |
| PPI                 | 3,890               | 76,584              | 50                  |
| Wikipedia           | 4,777               | 184,812             | 40                  |
**Experiment — Settings**

- **Multi-label Classification:**
  - **Label:**
    - Blogger classification
    - Biological state classification
    - POS tag classification
  - **Classifier:**
    - Logistic regression (L2)

- **Query-Based (Item) Recommendation:**
  - Recommending items to user based on user’s accessed items
  - Train:Test = 8:2

### Classification

| Classification     | |V| | |E| | #(labels) |
|--------------------|---|---|---|---|---|
| BlogCatalog        | 10,312 | 333,983 | 39 |
| PPI                | 3,890 | 76,584 | 50 |
| Wikipedia          | 4,777 | 184,812 | 40 |

### Recommendation

<table>
<thead>
<tr>
<th>Recommendation</th>
<th>#(users)</th>
<th>#(items)</th>
<th>#(edges)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Movielens-1M</td>
<td>6,070</td>
<td>3,706</td>
<td>1,000,209</td>
</tr>
<tr>
<td>Movielens-10M</td>
<td>69,878</td>
<td>10,677</td>
<td>10,000,054</td>
</tr>
<tr>
<td>Last.fm-2k</td>
<td>1,892</td>
<td>92,800</td>
<td>92,834</td>
</tr>
</tbody>
</table>
Multi-label Classification:
- Label:
  - Blogger classification
  - Biological state classification
  - POS tag classification
- Classifier:
  - Logistic regression (L2)

Query-Based (Item) Recommendation:
- Recommending items to user based on user’s accessed items
- Train:Test = 8:2

Baseline:
- SkipGram Based Method
  - DeepWalk
  - node2vec
- Edge Sampling Based Method
  - LINE (2nd)
  - HPE
Experiment — Multi-Label Classification

<table>
<thead>
<tr>
<th></th>
<th>BlogCatalog</th>
<th>PPI (Homo Sapiens)</th>
<th>Wikipedia (POS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DeepWalk</td>
<td>0.230</td>
<td>0.179</td>
<td>0.137</td>
</tr>
<tr>
<td>Node2vec ((p, q))</td>
<td>0.251 (0.25, 0.25)</td>
<td>0.185 (4, 1)</td>
<td>0.115 (4, 0.5)</td>
</tr>
<tr>
<td>LINE</td>
<td>0.231</td>
<td>0.170</td>
<td></td>
</tr>
<tr>
<td>NavWalker-clossness(^{-1}) (Improvement)</td>
<td>0.265 (5.5%)</td>
<td>0.192 (3.7%)</td>
<td>0.156</td>
</tr>
</tbody>
</table>
## Experiment — Query-based Recommendation

### MovieLens1M

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>MAP@10</th>
<th>Recall@10</th>
</tr>
</thead>
<tbody>
<tr>
<td>LINE</td>
<td>0.303*</td>
<td>0.377*</td>
</tr>
<tr>
<td>DeepWalk</td>
<td>0.270*</td>
<td>0.333*</td>
</tr>
<tr>
<td>Node2vec ($p = 1, q = 0.5$)</td>
<td>0.269*</td>
<td>0.331*</td>
</tr>
<tr>
<td>HPE</td>
<td>0.218*</td>
<td>0.283*</td>
</tr>
<tr>
<td>NavWalker (closeness)</td>
<td>0.333</td>
<td>0.411*</td>
</tr>
<tr>
<td>NavWalker (closeness$^{-1}$)</td>
<td>0.338</td>
<td>0.417</td>
</tr>
</tbody>
</table>

### MovieLens10M

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>MAP@10</th>
<th>Recall@10</th>
</tr>
</thead>
<tbody>
<tr>
<td>LINE</td>
<td>0.308*</td>
<td>0.375*</td>
</tr>
<tr>
<td>DeepWalk</td>
<td>0.275*</td>
<td>0.334*</td>
</tr>
<tr>
<td>Node2vec</td>
<td>0.164*</td>
<td>0.202*</td>
</tr>
<tr>
<td>HPE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NavWalker (closeness)</td>
<td>0.281</td>
<td>0.348</td>
</tr>
<tr>
<td>NavWalker (closeness$^{-1}$)</td>
<td>0.315</td>
<td>0.391</td>
</tr>
</tbody>
</table>

### Last.fm

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>MAP@10</th>
<th>Recall@10</th>
</tr>
</thead>
<tbody>
<tr>
<td>LINE</td>
<td>0.270*</td>
<td>0.295*</td>
</tr>
<tr>
<td>DeepWalk</td>
<td>0.276*</td>
<td>0.300*</td>
</tr>
<tr>
<td>Node2vec ($p = 0.5, q = 4$)</td>
<td>0.283*</td>
<td>0.295*</td>
</tr>
<tr>
<td>HPE</td>
<td>0.319*</td>
<td>0.361*</td>
</tr>
<tr>
<td>NavWalker (closeness)</td>
<td>0.169</td>
<td>0.190</td>
</tr>
<tr>
<td>NavWalker (closeness$^{-1}$)</td>
<td>0.343</td>
<td>0.396</td>
</tr>
</tbody>
</table>
Take Home Message & Future Work

Take Home Messages

• Rooted based random walk sampling is suitable for learning network representation.

• The flexibility of proposed framework facilitates the learning of information-enhanced representation.

• Centrality information strengthens the quality of learned representations.

Future Works

• Exploring other auxiliary information.

• Investigation for information enhancement methods.
Any Questions?

I am Kwei-Herng Lai
Research Assistant@AS-CITI
Contact: khlai@citi.sinica.edu.tw

Thank You!